DOCUMENT RESUME

ED 049 282 TM 000 454

AUTHOR	Flynn, Timothy M.
TITLE	Convergent-Discriminant Validation of Behavioral Ratings.
PUB DATE	Feb 71
NOTE	12p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Behavior Rating Scales, Cognitive Development, Comparative Analysis, Correlation, *Factor Analysis, *Migrant Children, Observation, *Preschool Programs, Program Effectiveness, Self Control, *Student Behavior, Teacher Aides, Validity
IDENTIFIERS	PKS, *Pre Kindergarten Scales

ABSTRACT

To ascertain the effect of a preschool program on the behavior of migrant children the Pre-Kindergarten Scale (PKS) was completed by teachers and teacher aides. In order to analyze the resulting data it was necessary to construct and validate the trait measures obtained from the scale. A sample of the ratings on 144 migrant children, subjected to a principal components factor analysis, provided four factors: Cognitive skills, self control, relationship with achievement model, and dependency. Factor coefficients were obtained for the 25 items based on four factors and applied to ratings on another sample of 153 migrant children. Trait scores for each rater were correlated with those of the other raters for a total of three multi-rater, multi-trait matrices. When these matrices were analyzed, cognitive skills, self control, and relationships with achievement model showed a convincing degree of convergent-discriminant validity. Dependency showed slightly less convincing validity. The findings support the overall validity of the four traits and have implications for future analysis of traits obtained from behavioral ratings. (LR)

ED049282

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Convergent-Discriminant Validation

of

Behavioral Ratings

Paper Presented at AERA Convention

February 4, 1971

New York, New York

TM 000 454

by

Tim Flynn

Child and Family Department

Southern Illinois University

Carbondale, Illinois

Convergent-Discriminant Validation of Behavioral Ratings

In an evaluation of the effect of a preschool program on the behavior of migrant children (CA = 3-11 to 4-11) a behavior rating scale (Pre-Kindergarten Scale, PKS) was completed by each child's teacher and the two teacher aides. In order to analyze the data obtained through the use of the PKS, it was necessary to ascertain the behavioral attributes measured by the scale. Initially the items on the scale were grouped under five headings according to the explicit factors they supposedly were measuring. However after the initial analysis of the data, it was apparent that the five explicitly constructed factors did not possess sufficient independence to be considered separate factors.

In an attempt to overcome this weakness it was decided to use an orthogonal factor analysis to obtain independent factors. The difficulty with this approach is that there is no assurance that factors derived from this procedure will be obtainable on analysis of additional data. However, there was available a procedure to determine the generality of the factors obtained. It consisted of obtaining factor coefficients from one sample of ratings, applying these coefficients to another sample, and then analyzing the obtained results.

This analysis not only provided a method to establish the independence and validity of the traits, but also provided support for the value of factor coefficients as item weights. The analysis performed consisted of subjecting the factor scores obtained from

each of the three raters (the teacher and the two teacher aides) to Campbell and Fiske's (1959) convergent-discriminant validation procedure.

Description and Employment of the PKS

Since the teacher aides in the preschool program were selected from the adults migrant population they possessed limited educational backgrounds. This made it necessary to employ a scale that not only would require a minimum of time and effort but would only require a grade school reading level of the rater. The PKS used a multiple choice format to simplify the rating procedure by providing both the situation in which the child was observed along with four alternative behaviors designed to describe the child's possible range of behaviors in the given situation. Each item was provided with a "have not observed" response. This allowed the observer to avoid rating if she had not observed the child in the given situation. The four responses for each of the 25 items were weighted one through four with the raters marking their responses on an IBM answer sheet. The following item from the PKS illustrates the multiple choice format:

When emotionally or physically upset, this child:

1. turns to his teacher for comfort and reassurances
2. accepts and responds to unsolicited comfort and reassurances from the teacher
3. passively accepts unsolicited comfort and reassurances
4. resists teacher's attempts at comfort and reassurance
5. have not observed

The teacher and the aides were provided with training in use of the PKS prior to and during the preschool program. Conversations with the teachers during this training indicated that they and the aides possessed an adequate understanding of the PKS rating procedure.

Factor Analysis of Ratings

The initial analysis was performed on the ratings of 144 children by the teachers and the aides. Three separate analyses were performed for each factor analytic approach used: (1) teachers' ratings, (2) aides' ratings, and (3) teachers' and aides' ratings combined. Three estimates of communality were used: (1) ones, (2) largest absolute row values, and (3) squared multiple correlation. For these three communality estimates both orthogonal and oblique rotations were performed. However, no significant differences were ascernable between the different estimates of the communality or between the orthogonal and oblique rotations. Since the PKS contained 25 items it was estimated that each item would be responsible for at least four percent of the variance by chance alone. Thus, a factor that did not account for more than five percent of the variance was not rotated. According to this criterion four factors were obtained for both the teachers' and aides' ratings. Since these factors were nearly identical, the teachers' and aides' were combined to obtain the factor coefficients.

Because all estimates of communality used with both orthogonal and oblique rotations gave similar results, the data was analyzed using a principal components (one's in the diagonal) approach. This made possible the obtaining of the factor measurements directly

since the inverse of the factor matrix existed (Harmon 1969). Another advantage of this approach is that it still applies when only a few of the larger factor components are used.

The four factors were interpretable and were given the following labels of: (1) cognitive skills, (2) self control, (3) relationship with achievement model, and (4) dependency. Factor coefficients were obtained from a principal component, orthogonal analysis, using the teachers' and the two teacher aides' PKS ratings of 143 migrant children. The output data consisted of 25 sets of four factor coefficients; that is one set of four coefficients for each of the 25 items on the PKS.

Convergent-Discriminant Validation of the PKS

The 25 sets of factor coefficients which were obtained through the factor analytic procedure were applied to another sample consisting of 153 migrant children who had been rated by their teacher and two teacher aides. Therefore, each child received four factor scores from each of the three raters for a total of twelve factor scores.

The three sets of factor scores for each child made it possible to adapt Campbell and Fiske's (1959) convergent-discriminant validation procedure to validate the four factors. The different observers, the teacher and two aides, were considered as different methods of obtaining the same information. Comparing the correlations between the four factor scores of three different observers provided a method of evaluating both the divergent and convergent validity of the four traits.

The comparison of the ratings by different observers of

the same individual is commonly referred to as a measure of inter-rater reliability. According to Campbell and Fiske, (1959) reliability measures are obtained by use of two measures of the same trait obtained from maximally different methods. The present investigator contends that the teachers' and aides' ratings would not qualify as "maximally similar methods", since their observations are relatively independent with the only similarity being the scale on which they were recorded. Moreover, since independence is a matter of degree, it seems that the use of different observers would fall more within the validity portion rather than within the reliability portion of the continuum.

As there were three observers it was possible to construct three multi-trait multi-rater matrices. The matrices are presented in Table 1 and consist of the correlations between the factor ratings of the three observers.

To establish a minimal degree of validity by using Campbell and Fiske's procedure it was necessary to demonstrate that the correlations between two different observers (methods) measuring the same trait were higher than the correlations between the unlike traits which were measured by the same or different observers. This was established through the examination of the validity diagonal of the three correlation matrices. The validity diagonal contains the correlation coefficients of like traits rated by the different observers. The entries in the validity diagonal not only should be significantly different from zero, but should be sufficiently large to encourage further examination of validity.

Table 1

Multi-Trait Multi-Rater Correlation Matrices (N=153)

		Cognitive Skills (CS)	Self Control (SC)	Relationship with Achievement Model (RAM)	Dependency (D)
Matrix 1			Teacher		
	CS	.51	.00	.29	.19
	SC	-.03	.47	.01	-.04
aide	RAM	.38	.05	.53	.12
1	D	.16	.16	.20	.28
Matrix 2			Teacher		
	CS	.42	-.05	.36	.03
	SC	.01	.57	.13	.06
aide	RAM	.25	.06	.46	.01
2	D	.27	.03	.13	.28
Matrix 3			Aide 2		
	CS	.32	-.11	.26	.03
	SC	-.19	.44	.10	.17
aide	RAM	.08	.03	.53	.21
1	D	.07	-.09	.30	.33

Only if the validity diagonal is sufficiently elevated does it provide support for the existence of convergent validity.

To aid in analysis of the data from the three matrices in Table 1, the correlation coefficients were averaged to provide comparison of the convergent-discriminant validity of the four traits. The convergent validity coefficients on the three diagonals in Table 1 provide support for the convergent validity of the following three traits: (1) cognition, (2) self control, and (3) relationship with achievement model. These traits possessed an average (over the three matrices) convergent validity of .42, .47, and .50 respectively. The trait dependency, however, appeared relatively weak in convergent validity with an average (over the three matrices) coefficient of .30. A possible explanation for this weakness lies in the method used to measure dependency. The context in which this trait is measured consists of the child's dependency reactions to the specific observer rating the child. Since the child's dependency behavior is to a certain extent contingent upon the adult whom he is interacting with, this would tend to weaken the relationship between the measures of dependency obtained from different observers. On the whole, these findings provide support for the convergent validity of the four traits.

While support for the convergent validity of the four traits is provided, it would be of little value without discriminant validity. On examination of the three matrices the existence of discriminant validity is supported by the elevation of the validity diagonal coefficient over their respective row and column

coefficients not on the validity diagonals. The comparison between the average of the 12 validity coefficients of .43 with the average of .11 for the 36 coefficients not on the validity diagonal provides support for the overall discriminant validity of the four traits.

The discriminant validity of the individual traits was also supported by the separate analysis of the elevation of each trait over its corresponding off diagonal coefficients. The results of this analysis demonstrated that the trait dependency with its average convergents validity of .30, and with its correlation of .11 with unlike traits demonstrated the least convincing evidence of discriminant validity with only a .19 overall elevation over its corresponding off diagonal coefficients. On the other hand, the remaining three traits produced more convincing evidence of convergent-discriminant validity as cognitive skills, self control, and relationship with achievement model possessed an average correlation with unlike traits of .10, .02, and .16 respectively, while their average convergent validity was .42, .47, and .50 respectively. This comparison demonstrates the high overall discriminate validity of these three traits.

Conclusions

Examinations of the three matrices presented in Table 1 indicated that the four traits established through use of factor coefficients on the PKS fulfilled the convergent-discriminant requirement that ratings of the same trait by different raters correlate higher than ratings of different traits by the same or different raters. This finding has several implications for

future research. For example, it supported the use of these four traits in the study of achievement motivation in the migrant child (Flynn and Curtis, 1970). These findings provide additional support for the value of the multiple choice format in obtaining objective information from teachers' observations. Previous use of the format on a scale designed to be used by teachers in rating educable, mentally retarded and normal elementary students indicated that use of multiple choice ratings made it possible to discriminate between retarded and normal students with 84 percent accuracy (Flynn, 1970). Accordingly, this format should prove useful in other situations where behavioral ratings are presently used. Not only did the findings validate the traits obtained from the PKS, they also demonstrated the value of factor analysis as a method for obtaining item weights. The fact that the coefficient obtained from a previous factor analysis generalized and provided relatively independent and valid trait measures from behavioral ratings demonstrated the potential of this technique.

If the present adaptation of Campbell and Fiske's (1959) convergent-discriminant validation procedure is not considered a true measure of validity, it still has value in that it provides a valuable method for illustrating the independence of trait measures obtained through behavioral ratings. Demonstrations such as this could also be employed to assess the relative independence of traits measured by other existing behavioral rating devices.

Summary

The purpose of the study was to construct and validate the

trait measures obtained from the Pre-Kindergarten Scale (PKS). The PKS was designed to be used by teachers and teacher aides in rating the behavior of migrant children (CA 3-11 to 4-11) who were attending a preschool program. A sample of the ratings of 144 children rated by their teachers and the two aides was subjected to a principal components factor analysis. Four interpretable factors were derived: (1) cognitive skills, (2) self control, (3) relationships with achievement model, and (4) dependency. Factor coefficients for the 25 items based on four factors were obtained and applied to another sample of teachers' and aides' ratings of 153 migrant children. Thus for each of the three raters there were four trait scores. The four trait scores for each of the raters were correlated with the other two raters for a total of three multi-rater multi-trait matrices. These three matrices were analyzed using Campbell and Fiske's (1959) convergent-discriminant validity procedure. Cognitive skills, self control, and relationships with achievement model possessed a convincing degree of convergent-discriminant validity while dependency possessed slightly less convincing evidence of convergent-discriminant validity. Not only did the findings provide support for the overall validity of the four traits, but they also had several implications for future analysis of traits obtained from behavioral ratings.

References

Campbell, D., and Fiske, D. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Flynn, T. M. Development of a multiple-choice behavioral observation scale. In Pilder (Ed.). Abstract One, Washington, D. C. American Educational Research Association, 1970.

Flynn, T. M., and Curtis, H. A. Traits related to achievement motivation in migrant children. Florida State University, Migrant Project Report, 1970.

Harmon, H. H. Modern Factor Analysis. University of Chicago Press, Chicago, 1969.